REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
1/20/98

3. REPORT TYPE AND DATES COVERED
Annual Technical 2/1/96 - 1/31/97

4. TITLE AND SUBTITLE
Training Spatial Knowledge Acquisition Using Virtual Environments

5. FUNDING NUMBERS
N00014-96-1-0379

6. AUTHOR(S)
Nathaniel I. Durlach

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Research Laboratory of Electronics
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139

19980126 135

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
Ballston Center Tower One
800 North Quincy Street
Arlington, VA 22217-5660

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
96PR03143-00

11. SUPPLEMENTARY NOTES
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The purpose of this study is to explore the potential benefits of using current VE technology for training spatial orientation and navigation skills.

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

FIRST ANNUAL REPORT FOR:

TRAINING SPATIAL KNOWLEDGE ACQUISITION USING VIRTUAL ENVIRONMENTS

(1 February 1996 to 31 January 1997)

NATHANIEL I. DURLACH, ET AL.,

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
RESEARCH LABORATORY OF ELECTRONICS
77 Massachusetts Avenue, Room 36-767
Cambridge, Massachusetts 02138
617-253-2534
DURLACH@CBGRLE.MIT.EDU

PREPARED FOR:
DR. T. ALLARD
COGNITIVE & NEURAL SCIENCES
OFFICE OF NAVAL RESEARCH (CODE 342)
800 N. QUINCY STREET
ARLINGTON, VA 22217-5660
703-696-4502 VOICE
703-696-8343 FAX

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) MIT/RLE/SEN/SPLOC	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Mass. Inst. of Tech. / RLE; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) 77 Mass. Ave #36-767 Cambridge, MA 02139-4307	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING / SPONSORING ORGANIZATION ONR/ Cognitive & Neural Sci.; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-96-1-0379 MIT Project 64333
8c. ADDRESS (City, State, and ZIP Code) 800 N. Quincy St. Code 342 Arlington VA 22217	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. / PROJECT NO / TASK NO / WORK UNIT ACCESSION NO

11. TITLE (Include Security Classification)
First Annual Report for:
Training Spatial Knowledge Acquisition Using Virtual Environments

12. PERSONAL AUTHOR(S)
Nathaniel I. Durlach

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical Final	FROM 2/1/96 TO 1/1/97	1/15/98	31

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The purpose of this study is to explore the potential benefits of using current VE technology for training spatial orientation and navigation skills.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code) / 22c. OFFICE SYMBOL

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

Acknowledgments

This work was performed under ONR Grant #N00014-96-1-0379 (MIT project #64333) for the Office of Naval Research. Nathaniel I. Durlach served as Principal Investigator. Dr. Terry Allard was the Navy Program Officer.

The following people contributed to the writing of this annual report: Dr. Thomas E.v. Wiegand, Nathaniel I. Durlach, Glen Koh, Maciej Stachowiak, & Bryan Kincy.

EXECUTIVE SUMMARY

This report summaries the work done by MIT Year 1 of the Training Spatial Knowledge Acquisition Using Virtual Environments program (1 Feb. 1996 through 31 Jan. 1997).

During this first year, we have pursued a number of parallel efforts in order to begin collecting experimental data as quickly as possible. The major experiments will be carried out using a large-scale photorealistic model of a 89,000 square foot warehouse. This venue was selected (in consultation with NPS) after examining a number of options on and near the MIT campus. We have designed a series of experiments to utilize this large-scale warehouse VE when the associated modeling work is completed. In order to begin familiarizing ourselves with various experimental issues as soon as possible, we performed an auxiliary preliminary experiment during the first year, using the same hardware and software to create a virtual representation of our seventh floor laboratory space. This first study explored the extent to which different types of VE presentations (head mounted display vs flat panel, use of a scaled down rotatable 3D model, etc.) are an effective means of conveying configurational knowledge of a specific real space. In addition, our project has supported ancillary work by a number of undergraduate students (UROPS). These projects included: (1) a study of VE construction for use in wayfaring exercises utilizing the VE action game Quake and associated shareware editors; (2) the development of a room-scanning robot for the purpose of systematically collecting texture map data from large and complex venues; and (3) the development of a "slippery-floor" mobility interface to enable more natural locomotion in virtual spaces.

FIRST ANNUAL REPORT FOR: TRAINING SPATIAL KNOWLEDGE ACQUISITION USING VIRTUAL ENVIRONMENTS

1. INTRODUCTION

The overall objective of this work, which is being conducted in collaboration with Rudy Darken at the Naval Postgraduate School, is to apply state-of-the-art virtual reality simulation technologies to the investigation of issues in spatial knowledge acquisition and training. Virtual environment training provides a number of advantages over traditional training methods and it is expected that such training will translate into measurable real-world gains in performance. The first set of studies are intended to explore the extent to which virtual environments are an effective means of conveying spatial knowledge of a specific real space. The second set, which will not be initiated until later on in the program, will be concerned with the training of spatial knowledge and navigational skills in general and with the ability to transform spatial information presented in one form to information presented in other forms. Throughout all of our work, we will explore the role of various environmental features and stimulus characteristics in navigation and wayfinding so that we determine which components of virtual-environment fidelity can be degraded without causing substantial reductions in training transfer. In this manner, we hope to develop systems that are optimally cost effective. To the extent possible, we will also attempt to shape our research to the needs of special communities (such as special operations groups) for whom detailed accurate spatial knowledge of spaces that are not accessible for training is of great practical importance. Because the work involves both basic and applied research, and is also highly interdisciplinary (it involves cognitive science, perceptual psychology, virtual-environment design, real-world training methods, etc.), it obviously requires close collaboration among a wide variety of individuals. General background on this program can be found in our initial proposal.

The first year of work on this project has been characterized by concern with the acquisition of spatial knowledge for specific spaces and with the development of facilities and procedures for conducting experimental work. Furthermore, attention has been focused on the acquisition of configurational knowledge as opposed to route knowledge. Whereas route knowledge is characterized by a navigator's ability to move from one position to another along a specific route, configurational knowledge is characterized by information on the relationships between different locations and the overall structure of the space, as well as by the ability to assume exocentric viewpoints. A navigator with strong configurational knowledge can find optimum routes between two points in the environment, draw maps of the environment, and estimate the direction and distance to a particular landmark from an arbitrary position in the environment. We have focused on configurational knowledge rather than route knowledge because it is fundamentally more important and because the use of virtual environments for providing route knowledge has been more extensively studied (e.g., see Witmer et al., 1995).

2. PROGRESS REPORT

Our work during the past year can be summarized under the following headings: (1) Experimental Design, (2) VE Implementation, (3) Support Efforts, and (4) Preliminary Experiments.

2.1. Experimental Design

The discussion in this section represents our initial plans for experimental design, i.e., plans made prior to conducting the preliminary experiments described in Sec. 4. During the second year, we will modify these plans in accordance with what we learn from the preliminary experiments.

2.1.1. Task Selection..

Configurational knowledge will be tested using two methods, each of which controls for external factors such as verbal descriptive abilities: a *pointing* task in which the subject is brought to a position A and asked to estimate direction and distance to a number of other positions B in the venue, and a *wayfaring* task in which the subject is asked to proceed from location A to location B via some route.

The *pointing* task measures the subject's ability to reproduce a 3D replica of the space by having him point and estimate the distance to landmark B from position A (which may not be the subject's current position). Theoretically, a subject who is able to identify the locations of all points relative to his own or relative to each other has a perfect spatial knowledge of the venue. The accuracy (in azimuth and distance) of the subject's responses can be used to evaluate the degree of knowledge of the space.

The *wayfaring* task involves having the subject move between two points in the venue via a route with characteristics described by the interviewer at the time of testing. The subject may be asked to find the shortest path between A and B or a path via a number of waypoints. The goal of this test is to assess the subject's spatial and navigational knowledge from an egocentric perspective. Because the test paths are not known to the subject during the practice phase of the experiment, the subject is forced to piece together observations of the venue derived from information obtained during the practice phase —it is not possible for the subject to memorize the specific routes to be tested. The subject's performance in the wayfaring task will be assessed using a number of variables, including: (1) time taken to complete the task, (2) number of wrong turns, (3) distance travelled by the subject, and (4) ability to complete the task.

2.1.2. Venue Selection.

The selection of an experimental site is clearly of great importance to continued success of the project. The site must be of sufficient complexity to reveal differences in the experimental conditions in current and subsequent studies. Because creating a virtual environment is time-consuming, it is important to choose a site which will be available and accessible for a period of time and also to create a model with sufficient detail to be useful for further experiments.

The benchmark for the virtual environment model is to produce the most accurate model and to create the deepest sense of immersion possible within the limits of VE modeling. Starting with a high-resolution model, it is possible to degrade the model in different respects as experimental conditions. Starting with a specific, less detailed, model would require the creation of a new model for each new VE feature introduced.

The goal, then, has been to select a venue with substantial complexity and to create a virtual model with sufficient detail that the VE model and real-world venue pair can be used for the next several years as an experimental system without complex modifications of the system (thereby amortizing the initial high modeling time-investment over a span of many experiments).

The venue chosen for modeling was MIT Building NW30. This building is a "dead storage" warehouse and was chosen because of its relative complexity, availability, and access for experimentation with large numbers of subjects. NW30 is on the MIT campus and approximately a 15 minute walk from our main laboratory at RLE (where the VE equipment is located). An additional appealing factor supporting the use of NW30 is that it is generally not visited by students, and hence does not rule out students as one of our classes of experimental subjects.

The model comprises four of the five levels of NW30. The basement of NW30 was not included in the model because the four above-ground levels provide sufficient complexity for the study. The NW30 building spans approximately 89,695 square feet. As indicated in Figure 1, a large number of loops and decision points are available. Loops are conduits which connect to themselves —if a subject follows a loop, the subject will eventually return to the starting point of the loop. Decision points are defined as branches in a path (or conduit) with two or more possible choices of direction. The complexity of a given venue is indicated by the number of decision points and loops in that venue. In Table 1 we summarize the single-floor-only (i.e, 2D) loops and decision points that are easily usable in the chosen venue.

Figure 1 Floor Plan of MIT Building NW30.

Building Level	Decision Points (not including stairwells)	Loops (2D only)
Floor 1	12	3
Floor 2	4	0
Floor 3	6	1
Floor 4	1	0

Table 1 Single-floor Decision Points and Loops in NW30.

Note that the selected venue contains a number of stairwells, and that each stairwell is also a decision point. The effective complexity of the building is thus multiplied because of the many loops which exist once the stairwells and the various 3D paths are taken into consideration. Additionally, the building contains an elevator and a mezzanine level which is generally not accessible; these features can be pressed into service if future experiments require the added complexity.

The selected building and its associated virtual model will allow for many subsequent experiments because of its complexity and multi-level nature. Complexity can be arbitrarily increased by selecting paths with multi-story waypoints. Conversely, it is possible to use the virtual model for a less complex navigational study simply by limiting the subject to one floor.

2.1.3.Training Conditions.

There are four training conditions to be used in connection with the main venue (a similar protocol, to be described in more detail in section 2.4., is used in the prototype experiment conducted in the main lab in MIT Building 36):

1. Traditional training using maps only.
2. Exploration of venue with maps.
3. Exploration of virtual reality representation of venue with maps.
4. Exploration of virtual reality representation of venue with maps and virtual reality training aids.

Subjects are allocated evenly to these four groups. Subjects have no prior knowledge or experience with the experimental venue. The experiment proceeds in two phases: the initial training phase utilizing one of the above four treatments, and a testing phase, elaborated further below.

The training phase allows subjects to acquire some form of knowledge about the venue. For each of the training conditions, the subjects are given a fixed amount of time during the training to digest and acquire some knowledge of the space. All subjects are given 20 minutes of training and are encouraged to spend the entire allotted time learning the environment.

Training involves giving the subject a map of the venue with certain landmarks and locations highlighted. All four groups of subjects receive identical maps as a learning tool. The subjects are notified that the highlighted points on the map are of particular importance in the task. Subjects in training conditions 2, 3, and 4 are allowed to explore some representation of the space charted by the map (the actual building for condition 2, and the virtual model for conditions 3 and 4). Subjects are given free exploration of the venue in both the virtual environment and the real world conditions with no input from the experimenter, aside from the suggestion that the subject visit all the highlighted locations. The only constraint on exploration is the venue itself (available conduits in both the real and virtual environments and the 20 minute time limit).

Subjects in treatments 3 and 4 are exposed to the virtual environment training system. They are given as much time as is necessary to familiarize themselves with the VE and the interface equipment. The subjects in these two groups are allowed to gain knowledge of the system and its navigational tools by exploring an unrelated "sample" space in the VE (consisting of a model of a part of our main laboratory). In this way, the subject will be able to learn to be comfortable operating in the VE, and will gain familiarity with the system without gaining extra time to learn about the experimental venue.

Two forms of the VE are available, assigned to the two VE training groups:

1. The first VE is the standard high-resolution model.
2. The second VE is enhanced to allow the subject to select "translucent walls on demand" in effect giving the subject a sort of "x-ray" vision.

Subjects in all four groups are given 20 minutes to explore the environment. At the end of the 20 minutes, subjects proceed to the actual venue for the testing phase.

2.1.4. Testing Procedures.

Subjects from the four treatment groups are tested in the actual venue. All subjects are presented with the same set of tasks; in both the pointing and the navigation tasks, the subjects are asked to identify the same locations and navigate the same routes.

The testing phase consists of six tasks: three pointing tasks and three navigational tasks. The tasks are interleaved, with subjects alternating between the navigational and the pointing tasks. Upon entering the experiment site, subjects are asked to walk to a position A from the entrance of the warehouse using the shortest possible route. At this target point, the subject is asked to point and estimate the distance to a location B. This process is repeated twice more, with the subject moving from the current position to a new position through a route described by the experimenter. After this cycle of three sets of two tests, the testing phase for that particular subject is complete.

The three pointing tests require the subject to:

1. Point at locations on the same level.
2. Point to locations on a different level.
3. Point at a location at an initial point displaced from his own. The subject will be asked to place himself in the frame of reference of a remote location and point toward a different remote location.

The wayfaring tasks will consist of:

1. Finding the shortest path to a given location.
2. Finding a path via a number of waypoints.
3. Finding a path to a location when an unexpected barrier is introduced.

Because the target points are identified in the real world venue, the subject should always be able to eventually reach the next location as identified by the interviewer. If the subject can not do so within a reasonable amount of time or seems lost (as determined by the experimenter), the subject is brought to the next location and asked to complete the pointing task at that location. Mistakes made by the subject are not identified by the interviewer to the subject. In tasks in which the subjects are expected to traverse a route via some set of waypoints, missing a waypoint should be obvious to the subject as the waypoints are easily identifiable locations or objects (water-fountains, fire extinguishers, etc.).

All subjects are evaluated with the same criteria independent of the training method. The tabulated data of Table 2 are then analyzed to determine the relationships between task performance and training method.

Navigational Task	Time Taken	Distance Travelled	# of Wrong Turns	# of Missed Waypoints
Wayfinding 1				
Wayfinding 2				
Wayfinding 3				
Pointing Task	**Azimuth**	**Distance**		
Pointing 1				
Pointing 2				
Pointing 3				

Table 2 Test-phase Measurements for each subject.

2.2. VE Implementation

The choice of hardware with which to create the VE was largely pre-determined by the availability of existing equipment in the lab. The choice of simulation software (to render the building model) was made to achieve parity with what was being used at NPS. Although much of the existing hardware and software was previously used in other configurations in earlier projects, integrating the separate components (i.e. trackers, HMD, motion control hardware) with the new simulation and rendering software was nevertheless a non-trivial task. In this section we describe some of the issues involved in integrating the available components to provide a "walk-through" simulation that allows a subject to look around and have his point of view in the simulation follow the angle of his head. We also describe the implementation of a prototype locomotion interface which allows for two different modes of control (forward/back and left/right turn), and we describe how we implemented a workable collision detection algorithm to prevent the subject from walking through walls and furniture.

2.2.1. Software summary.

The visual simulation software suite which evolved from Steve Lakowske's work on the NASA/Army Cockpit Display Editor (CDE), has become commercially available through the software company Coryphaeus, and is in common use in government labs to create virtual environments. In particular, our colleagues at NPS have used EasyScene and Designer'sWorkbench, as well as other Coryphaeus modules, to good effect in previous VE-oriented projects.

Our simulations were developed using Designers Workbench (DWB), a software tool which facilitates the creation of 3D models. The EasyScene module renders the virtual world in real-time, and includes provisions for tracking the subjects' actions, including the distance travelled, time spent moving, viewing direction, etc., which enables the subjects' actions in the virtual world to be characterized.

2.2.2. Hardware summary.

The simulated environment itself was developed using a Silicon Graphics workstation running the DWB software. The real-time VE rendering is performed on a 2 processor SGI Onyx with Reality Engine II graphics. A stereoscopic head-mounted display (HMD) is used as the subjects' visual interface to the virtual environment during the "immersion" training conditions. The HMD is keyed to the direction of the subject's head, as tracked by a Polhemus 6-DOF position/rotation sensor. Control devices for the subject include:

1. Mouse and Keyboard
2. Serial Joystick
3. Future locomotion devices: Finger Walker, Slippery Floor, Magic Wand

2.2.3. Integration of Hardware and Software: Some Issues.

The EasyScene package provides facilities for real-time rendering of the scene, including provisions for object behavior and interactivity through the use of external application programs. Thus the best approach to integrating new features appeared to be to write a program that uses the available interprocess communication channels to control the simulation program itself. Using this approach, timing issues would not be as critical, since the process receiving the input data and transforming it into movement and orientation could run completely asynchronously from the simulation. However, we soon discovered that this interface alone was not enough to support collision detection, and some additional code had to be linked with the simulation itself to support this feature.

2.2.3.1. *Implementation of serial input devices.*

Implementing the head tracking control system required attention to several details. Interfaces to the hardware devices had to be implemented. This required handling of the RS-232C protocol over a serial line, as well as handling of the device-specific protocols for the joystick and the tracker. One module implemented the RS-232C serial interface. Under the workstation's operating system, serial channels are typically expected to be interactive terminals, so much preprocessing of the input stream is done by the operating system itself by default. However, the devices used are not much like a terminal, and essentially none of this pre-processing makes sense, since they send a raw byte-stream with no desire for line buffering or control-flow protocols. Thus, the biggest challenge in using the devices correctly was turning off all of the terminal-related overhead. Another key issue was correctly setting the speed and other communication parameters. However, once these issues were dealt with, the rest of the serial module was quite straightforward.

The principal input hardware was a head mounted Polhemus Fastrak rotation/position sensor and a processing unit that interprets data from the sensor and sends it to the workstation via an RS-232 serial line. It was clear that the control process should listen to the serial port in raw mode, and as position/orientation data becomes available, translate it into a position and orientation in model space, and finally send it to the simulation process as described above.

One extension that was considered was to use a second tracking sensor connected to a wand to indicate desired directions of motion via gestures with the wand. However, this idea was discarded as unnecessarily complex, and instead a simple serial joystick provided the walking interface. Eventually, some sort of walkable surface sensor is planned, but this interface has not yet been completed.

The joystick hardware which provides the walking interface connects to a serial port and provides X-Y position information, and an additional coordinate derived from a knob on the side of the joystick. Several modes for use of the joystick are plausible interfaces for walking, and the decision on the actual walking interface was deferred until the implementation stage.

2.2.3.2. *Communication with real-time rendering software.*

The process that interfaces with the devices and calculates position data communicates with the rendering system through shared memory. This allows high bandwidth communication but allows the application to run as a separate process from the rendering engine to avoid interfering with the rendering pipeline. Very few problems resulted from using the shared memory interface, except on rare occasions when needed functionality could not be accessed through it.

2.2.3.3. *Coordinate transform of input device data.*

In addition to interfacing with existing hardware and software, a number of other issues had to be considered. First, coordinates from the head tracker must be translated into a coordinate system suitable for use with the simulation, and similarly for joystick coordinates. These two sources of input must also be combined in a coherent way.

The Polhemus sends its position and orientation relative to some initial state. However, these coordinates cannot be used directly to control the simulation. They must be mapped relative to the current camera position and orientation in the virtual space, and some accounting must be made for the fact that the sensor is not mounted between the eyes but rather on top of the head. This is simply a matter of applying suitable transforms to the coordinates received. However, it is also necessary to initially calibrate the sensor. These issues also apply to the joystick except that the initial state is known ahead of time so the calibration step is unnecessary.

No single interface for combining the joystick and tracker coordinates met all of the projects needs. Thus, three different interfaces were implemented. The first is a "wheelchair" style arrangement with independent head motion. A current facing of the wheelchair is maintained internally by the program, and moving the joystick sideways rotates the wheelchair in place. Pushing the joystick forward or back moves the subject along the current wheelchair facing. The position and orientation of the viewer's head is taken relative to that of the wheelchair. Thus, it is possible to look sideways and move forwards, and see the scenery on one side pass by.

This first interface was considered potentially unsuitable for acquisition of spatial knowledge. Thus, two more were implemented which did not maintain a separate orientation for the point of view and the direction of forward motion. The second interface completely ignores the sideways motion of the joystick, and allows forward or backward motion to move the subject along the current point of view (with the z coordinate removed to avoid up and down motion). The third interface mode is essentially the same as the second except that moving the joystick sideways moves the subject sideways relative to the current point of view, allowing side-stepping.

2.2.3.4. *Implementation of collision detection.*

To achieve some degree of physicality and interaction in the VE, it is necessary to implement collision detection with respect to the walls (and other objects) during a walk-through of the scene, i.e., the user must not be permitted to pass through solid objects in the model. The EasyScene software already provides the means to do this collision detection, but it must be suitably enabled, and considered when controlling the motion of the point of view. Unfortunately, full collision-detection functionality is not available through the shared memory interface. Since the existing architecture was already working smoothly, a plan was devised that would allow it to work with only minimal modifications to the rendering engine. Essentially, a function in the simulation receives collision data, however, it communicates it through a separate IPC channel to the controlling process, which then handles collisions when they occur. The controlling process is able to directly enable collision detection and set collision geometries, but must resort to this back channel to be informed when collisions occur.

The virtual environment to be used includes stairs. This presents a special difficulty, since stairs must not be detected as a source of collision, but rather should gradually elevate the position of the point of view, while allowing forward motion to continue. Since the simulation should work with minimal foreknowledge of the target environment, and since the simulation does not directly have data on all objects, solving this problem is difficult and work on this aspect was deferred until a later time.

2.3. Support Efforts

2.3.1. Texture Acquisition System

Texture maps must be added to the DWB model in order to create a photorealistic virtual representation of the real-world venue. Collecting texture maps on a large scale is non-trivial, and manual methods of collecting and digitizing photographic images would become extremely time-consuming in such a large venue. Some form of automation must be used to facilitate the gathering of texture data.

One approach that is being successfully employed at several labs involves collecting image sequences along relatively unconstrained paths with some means of recording the camera position and orientation for each image. Advanced image interpolation algorithms can then be used to appropriately morph separate scenes and extract geometries in order to make an automatic record of the 3D space. Obviously, this approach represents a formidable research effort in itself, and a simpler alternative appropriate to our current project was sought.

2.3.1.1. Textured conduits.

One goal in selecting a venue to be modeled is that it should be visually complex: it should have many landmarks or clutter. The warehouse that we are basing the main model on fulfills this criterion, but also raises the possibility that the complexity will be too great to model. However, because we are primarily concerned with navigation along clear paths, (and there are many relatively clear corridors/hallways throughout the building) the "clutter" can be limited to the perimeter of the path, and confined to the various rooms or areas that open up off of the corridors.

With this conceptualization, much of the space can then be modeled as a "network of tubes (of rectangular cross-section)" which represent the corridors. "Ribbons of texture" which convey views of the visual clutter can then be texture-mapped on the inside surfaces of these tubes, or rather, "ductwork." This large human sized ductwork is analogous to the Habitrail system used to create complex 3D spaces for pet hamsters to play in.

By basing our VR model-building efforts on this "ducts and ribbons of texture" approach, we can preserve the visual richness without actually modeling an overabundance of individual objects.

2.3.1.2. Telecentric room scanner.

Collecting the required texture data requires only that the recorder follow the "pedestrian" paths through the piles of clutter in the warehouse, just as a person would. We are building such a device to perform this task. The room scanner consists of a Quick-cam video camera connected to an Apple Macintosh computer and placed on a rolling cart with a stepper motor attached to its wheel. The computer controls the stepper motor, which advances the device to the next incremental position, then grabs the next frame, etc. making a continuous run to collect the series of one-pixel-wide vertical strips, which are stored as a spatially-addressable QuickTime movie file.

The image and position data (derived from the stepper motor) are stored in a multi-track QuickTime file as a sequence of 1-by-640-pixel frames. The resulting file, or "ribbon of texture," is addressable from arbitrary start and stop locations. Since we also include location information from the floor plan, we can automatically extract the textures from the QuickTime files and apply the strips of texture on the conduits defined in the DWB model.

By collecting one-pixel-wide-frames at a time we avoid point-of-view perspective problems when assembling the multiple images. Using this method, the camera is always perpendicular to the surface it is viewing. (Actually, this is only true in the horizontal direction. The horizontal scan does nothing to eliminate the implicit vertical point of view, but by keeping the camera at eye level the perspective will be appropriate for a subject walking through the space.) Thus, we avoid the need for sophisticated image morphing algorithms to make multiple images fit together.

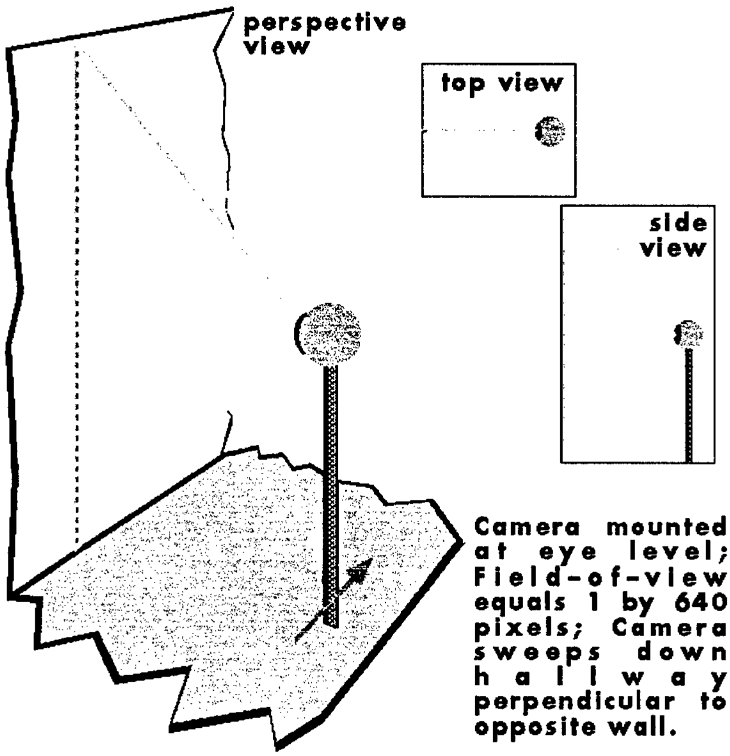

Figure 2 Room Scanner Conceptual Plan.

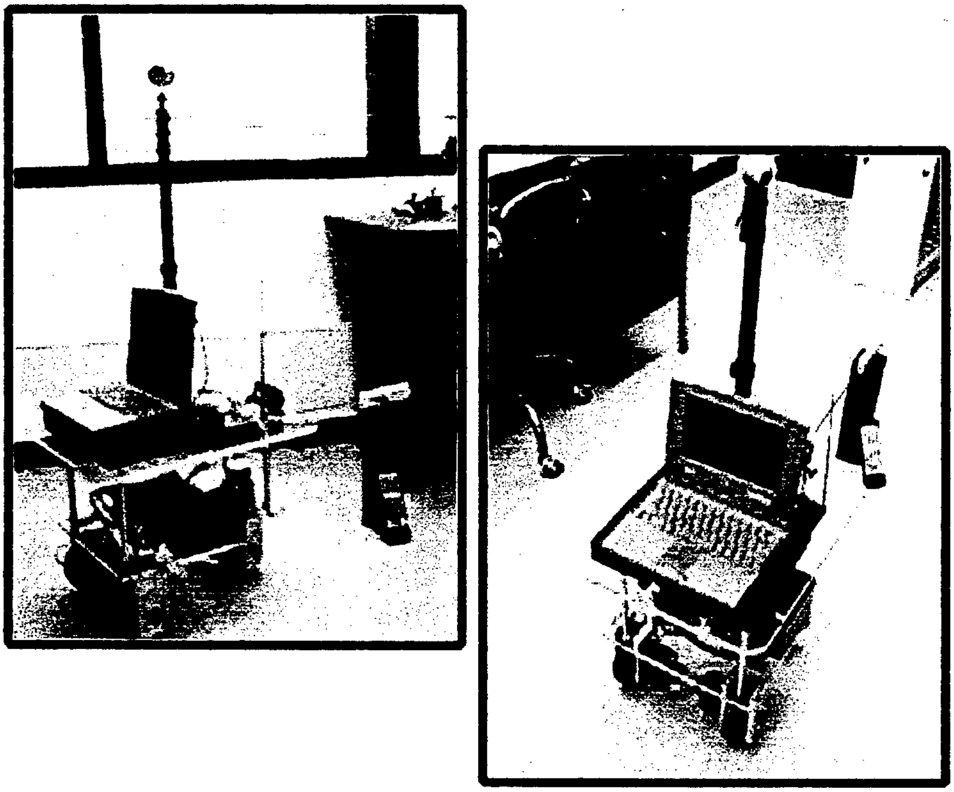

Figure 3 Room Scanner Prototype.

2.3.2.Slip-floor Mobility Interface.

An important aspect of the VE training experience involves providing a realistic or at least plausible means of moving about within the virtual world. The choice of motion control interface has obvious impact upon the general ergonomics and usability of the training system, but the choice of interface may also have less obvious but critically important interaction with the effectiveness of the system as a means of imparting spatial knowledge. One example of such an interaction which we would like to explore involves the expenditure of energy in moving from place to place as one does when walking.

The proposed "slip floor" consists of a low-friction foot-floor interface, some guying arrangement with slip bearing to allow azimuthal rotation, and some method for monitoring the position of the feet as they move in a walking motion. Through the use of appropriate pattern recognition algorithms, different foot motions could be resolved into motion control signals for use in controlling the VE simulation.

We have begun to explore the technical trade-offs of various sensing systems, and have sketched out a number of ideas which utilize the walking metaphor, including a system which is based on "finger-walking" and an electric-field sensing technology that may be applicable to large and small-scale walking interfaces. Results of this supplementary research should be available during the second year of the project.

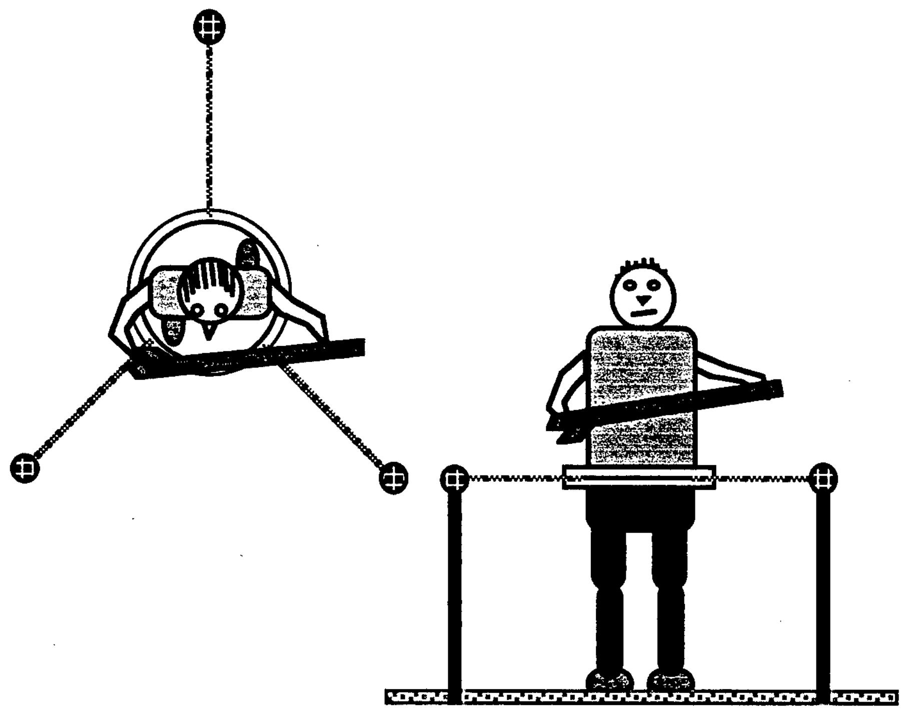

Figure 4 Slip-Floor Interface Conceptual Diagram.

2.3.3. *Low-end VE Feasibility Study.*

This side-project explores a low-cost approach to creating a VE for training which exploits a readily available shareware-based game development system running on a generic PC. One precedent for this approach is Barnett and Snyder's use of Doom for battlefield scenario simulation for the Marine Corps. For our simulation, we are using an improved version of Doom, called Quake, which includes a software-based 3D graphics engine. Using a shareware editor, one is able to construct virtual worlds to which run on the standard Quake application. These worlds, may include localized sound, 360 degree freedom of movement, collision detection, and interaction via projectiles.

In order to use Quake as a learning tool, one must first be able to construct new worlds from scratch, import textures from the 3D space being studied, and provide realistic lighting that mimics the given effect. While the range of editors that can perform these functions is small, there are new editors being released that provide a wider range of possibilities. At the current stage of development, the only setback has been the quality of the textures. Once better texture editors are produced, the possible applications of this approach will be expanded greatly.

Using a shareware editor, we have constructed a world that resembles the oddly shaped open space in our lab on the 7th floor of RLE. The world is lit similarly to the real space. Due to limitations in the current editor, textures on the walls had to be created manually and were not based on photographs of the real space. However when editors with improved capabilities become available, the application of textures from the real space will enhance the resemblance of the two worlds. Some of the more difficult objects to represent are the furniture. Because the level of detail the editors use, the furniture in the room must is crudely rendered, although an untrained eye should be able to determine what the images represent.

While the floor plan of the current 3D world is architecturally correct, the problems alluded to above need to be overcome before such a system could be used for experiments which require a close match to the visual clutter and surface qualities of some real-world venue. However, in cases where a relatively abstract representation of an architectural space is all that is required, the extremely low cost of such an approach is a strong vote in its favor.

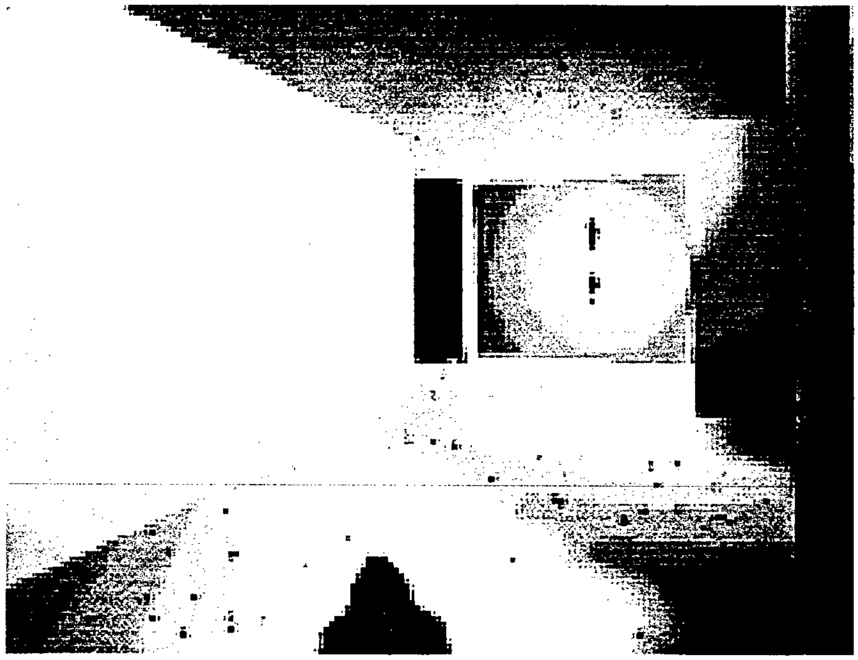

Figure 5 Quake Scene of 7th floor.

2.4. Preliminary Experiments

Conducting preliminary experiments was motivated by a number of factors. First, we thought the experience would help us refine the design of the main experiment. Second, conducting such experiments would enable us to test our facilities and processes for VE model building. Third, if we used a test venue distinct from the warehouse venue described in Sec. 2.1.2, then the test venue used in these preliminary experiments could serve as the familiarization venue used for the main experiments (See Sec. 2.1.3).

2.4.1. Venue.

The venue selected for the preliminary experiment consisted of a portion of the 7th floor of MIT building 36 (the same venue as that used for the low-end feasbility study mentioned in Sec. 2.3.3). This venue was readily accessible and located near to our VE laboratory facilities. Within the 7th floor area, most private offices were excluded; VE presentation and subject exploration were confined to public and certain laboratory spaces.

2.4.2. *VE Implementation.*

Implementation of this "sample" VE uses the same hardware/software facilities as those to be used for the large-scale warehouse. The issues covered in section 2.2. were largely worked out as a result of putting together this "sample" simulation. At the current time, we are involved in assessing and fine-tuning subtle but important aspects of the VE presentation, such as apparent discrepancies of scale occurring within the model, the creation of a natural-appearing field-of-view in the HMD, adjustment of light sources, etc. We are also resolving some issues related to collisions with objects in the virtual environment. As is readily apparent in comparing the photograph of one view of the real space (Figure 6) to a roughly corresponding screen-dump from the HMD view of the simulation (Figure 7), we have not yet achieved a perfect correspondence in many domains. A topic for future research will be to determine which of these presentation domains have the greatest impact on training transfer. However, at present, we are must perform this fine-tuning in an *ad hoc* manner.

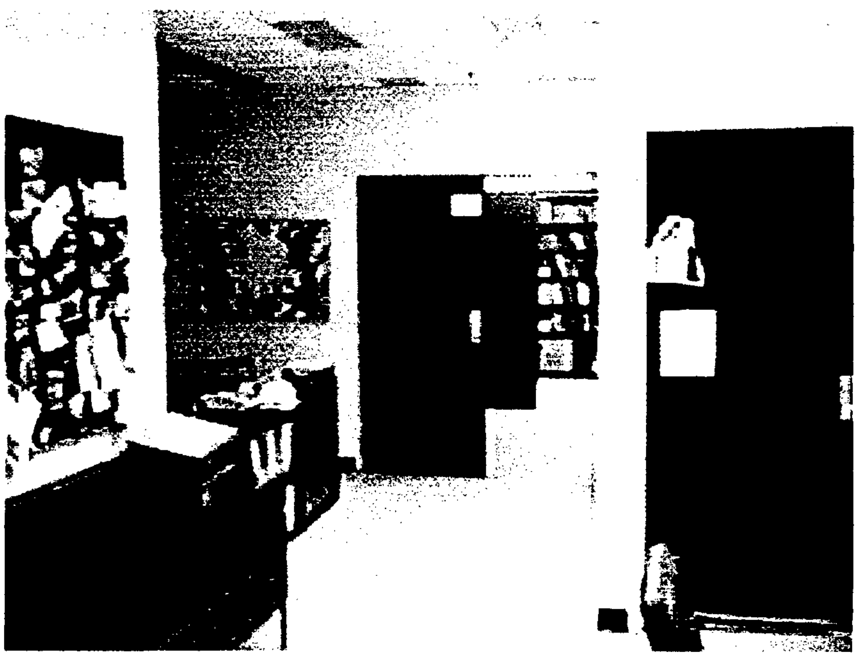

Figure 6 Veridical view of MIT 36-767.

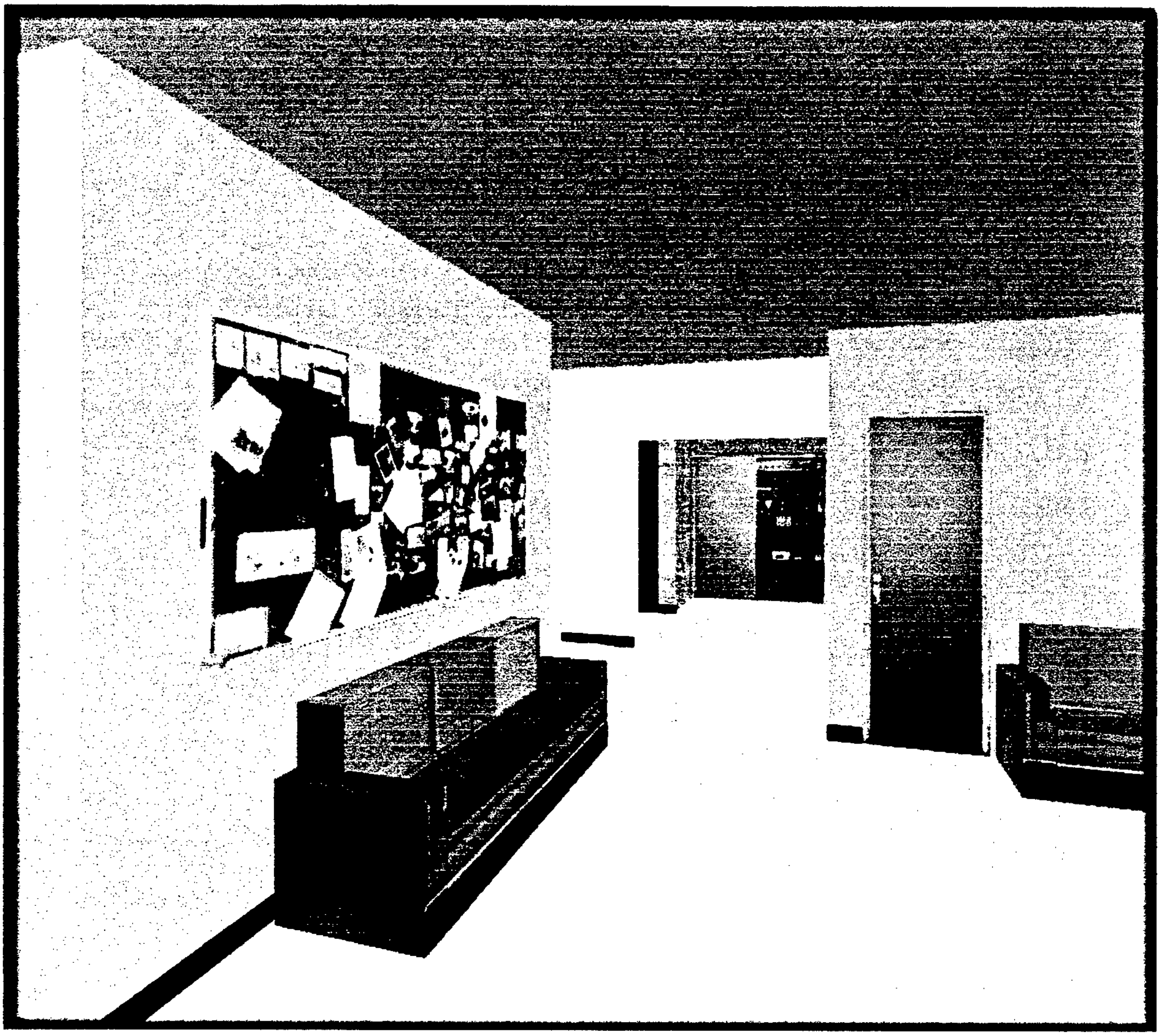

Figure 7 HMD view of reception area of MIT 36-767.

2.4.3. *Description of Preliminary Experiment.*

The test task employed was chosen to evaluate configurational knowledge, not route knowledge. In general, we regard configurational knowledge more important than route knowledge for two reasons. First, there exist tasks for which performance depends on configurational knowledge, but not on route knowledge (e.g., designing a ventilation system for a building, planning to demolish a building, "bugging" a building, etc.). Second, whereas configurational knowledge generally implies route knowledge (including knowledge of secondary routes when primary routes are blocked) route knowledge does not generally imply configurational knowledge.

For all training conditions, the subject, located at a given reference position within real space, was required to estimate the location of a given landmark within the real space by reporting the bearing of the landmark and the range of the landmark relative to the location of the reference. Thus, the type of knowledge being probed in these tests was not only configurational, but Euclidean-metric configurational (rather than, for example, topological or city-block-metric configurational).

Four training conditions were employed: RW (Real World); VE (Immersive VE); NVE (Non-Immersive VE); Mod (Model). For all these training conditions, the subjects were informed ahead of time about the task they would be asked to perform after the training was completed. Also, for all conditions, the training included 10 minutes of free exploration under the specified training condition. For the RW Condition, this was the only training provided. In the other three cases, an opportunity to become familiar with the technology was provided prior to the 10 minute training period in which the synthetic version of the 7th floor space was explored.

In the VE condition, subjects used a headmounted display (HMD) for visualization and a joystick for navigation. The NVE condition used the same equipment and procedures, except that a 21" monitor was used in place of the HMD. Subjects in each of the training groups for VE, NVE, and Mod, were familiarized with their training equipment by permitting them to explore a vastly simplified VE (unrelated to the test space). In the Mod condition, subjects were provided with a virtual miniature 3D model of the space that could be manipulated using a mouse and monitor. Subjects were able to rotate and view the model from any vantage point as well as zoom in closely to examine specific details. The graphical model of the space used for this condition was essentially identical to that used in the two virtual "walkthrough" conditions VE and NVE mentioned above. The only differences were the point of view change and removal of the ceiling so that the "internal space" could be seen from the outside. A sample view for the Mod condition is shown in Figure 8. Again, as with the VE and NVE conditions, for the Mod condition the subjects were provided with an opportunity to familiarize themselves with the system by exploring a model that was entirely distinct from the model to be used in the training experiment.

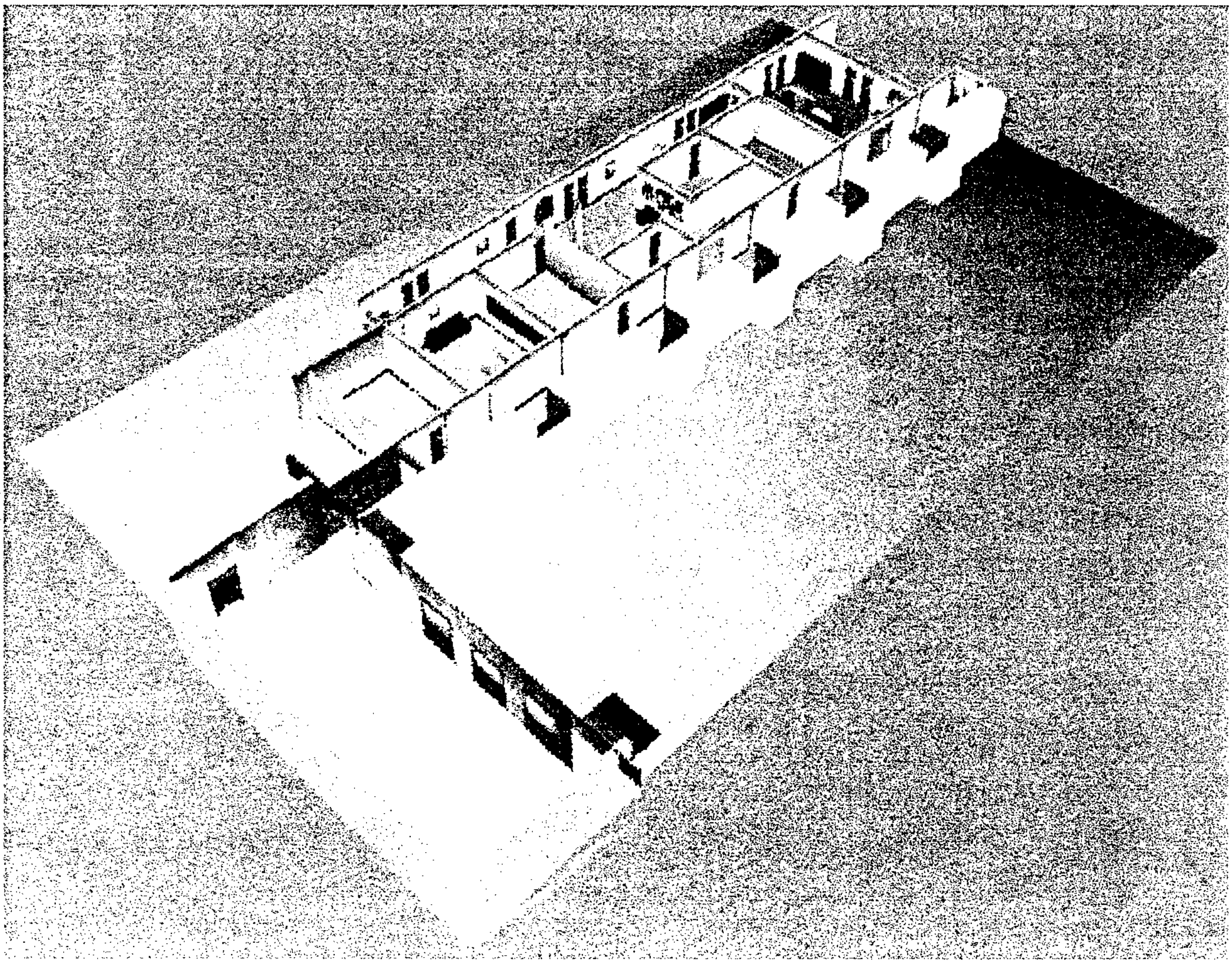

Figure 8 Stadium view of model of the 7th floor of RLE.

3. OVERVIEW OF WORK PLAN

During Y1 of this grant, we have developed the basic infrastructure for our experimental work. In particular, the hardware and software needed to create the envisioned VEs have been made operational, and a sample VE has been constructed for experimental use. Furthermore, we have conducted some preliminary experiments making use of this sample VE. Additional work on infrastructure has included initial development of (1) an automated "room scanner" for efficiently capturing textures for use in VE models and (2) a "slippery floor" mobility interface to serve as a more natural means for moving through VEs.

During Y2 of this grant, we intend to complete our preliminary study using the sample venue (data analysis and preparation of results for publication), make

During Y2 of this grant, we intend to complete our preliminary study using the sample venue (data analysis and preparation of results for publication), make significant progress on the modeling of the 89,000 square foot warehouse venue, refine our experimental design for the experiments to be conducted in this venue, and begin outlining our program for exploring the extent to which VE technology can be used to enhance spatial behavior *in general* (not only to comprehend and navigate through specific spaces).

Additional work on infrastructure during Y2 will include further development of the room scanner and the slippery-floor interface, and exploration of the feasibility of using common PC-based computers and inexpensive software to achieve increased cost-effectiveness in the VE training area.

During Y3 of this grant, we expect to complete work on the room scanner and slippery floor interface, conduct a series of experiments in the warehouse, begin to develop a PC-based virtual environment test bed, and initiate experimental work on training spatial behavior in general.

BIBLIOGRAPHY

Allen, G. L., Kirasic, K. C., Siegel, A. W., and Herman, J. F. (1979). Developmental Issues in Cognitive Mapping: The Selection and Utilization of Environmental Landmarks. Child Development, 50, 1062-1070.

Appleyard, D. (1973). Notes on Urban Perception and Knowledge. In R. M. Downs & D. Stea (Eds.), Image and Environment, (pp. 109-114). Chicago: Aldine Publishing Company.

Blades, M., and Spencer, C. (1988). The Development of 3- to 6-Year-Olds' Map Using Ability: The Relative Importance of Landmarks and Map Alignment. The Journal of Genetic Psychology, 151(2), 181-194.

Bower, G. H. (1970). Organizational Factors in Memory. Cognitive Psychology, 1, 18-46.

Broadbent, G., Bunt, R., and Llorens, T. (Eds.). (1980). Meaning and Behaviour in the Built Environment. New York: John Wiley & Sons, Ltd.

Darken, R. P. (1996). Wayfinding in Large-Scale Virtual Worlds. Unpublished Doctoral dissertation, The George Washington University.

Ekstrom, R. B., French, J. W., Harman, H. H., and Dermen, D. (1976). Manual for Kit of Factor-Referenced Cognitive Tests. Princeton, N.J.: Educational Testing Service.

Frankel, F., and Cole, M. (1971). Measures of Category Clustering in Free Recall. Psycho- logical Bulletin, 76(1), 39-44.

Hirose, M., Takahashi, K., Koshizuka, T., Morinobu, T., and Watanabe, Y. (1996). An Alternate Way to Generate Virtual Worlds: A study of image processing technology for synthetic sensations. PRESENCE 5 (1), 61-71.

Hirose, M., Takahashi, K., Koshizuka, T., and Watanabe, Y. (1994). A Study on Image Editing Technology for Synthetic Sensation. ICAT94, the Fourth International Conference on Artificial Reality and Tele-existence, July 14~15, 1994, Nikkei Hall, Otemachi, Tokyo. Koh, Glenn., MIT M.Eng. Thesis May 1997.

Krasner, L. (Ed.). (1980). Environmental Design and Human Behavior. New York: Pergamon Press.

Lynch, K. (1960). The Image of the City. Cambridge: MIT Press.

May, M., Péruch, P., and Savoyant, A. (1995). Navigating in a Virtual Environment With Map-Acquired Knowledge: Encoding and Alignment Effects. Ecological Psychology,7(1), 21.

McGee, M. G. (1979). Human Spatial Abilities: Psychometric Studies and Environmental, Genetic, Hormonal, and Neurological Influences. Psychological Bulletin, 86(5), 889-918.

Passini, R. (1984). Wayfinding in Architecture. New York: Van Nostrand Reinhold.

Péruch, P., Vercher, J., and Gauthier, G. M. (1995). Acquisition of Spatial Knowledge Through Visual Exploration of Simulated Environments. Ecological Psychology, 7(1), 1.

Piaget, J., and Inhelder, B. (1967). The Child's Conception of Space. New York: Norton.

Regian, J.W., Shebilske, W.L., and Monk, J.M. (1992). Virtual Reality: An Instructional Medium for Visual-Spatial Tasks. Journal of Communication 42(4), 136-149.

Riddell, R. (1997). Doom goes to War: The marines are looking for a few good games. Wired 4/97.

Roenker, D. L., Thompson, C. P., and Brown, S. C. (1971). Comparison of Measures for the Estimation of Clustering in Free Recall. Psychological Bulletin, 76(1), 45-48.

Seidel, A. D., and Danford, S. (Eds.). (1979). Environmental Design: Research, Theory, and Application. Washington, D.C.: Environmental Design Research Association.

Stoakley, R., Conway, M. J., and Pausch, R. (1995). Virtual Reality on a WIM: Interactive Worlds in Miniature. Proceedings of ACM SIGCHI 95, Denver, CO.

Thorndyke, P. W., and Goldin, S. E. (1983). Spatial Learning and Reasoning Skill. In H. L. Pick and L. P. Acredolo (Eds.), Spatial Orientation: Theory, Research, and Application, (pp. 195-217). New York: Plenum Press.

Thorndyke, P. W., and Hayes-Roth, B. (1982). Differences in Spatial Knowledge Acquired from Maps and Navigation. Cognitive Psychology, 14, 560-589.

Witmer, Bob G., Bailey, John H., and Knerr, Bruce W. (1995). Training Dismounted Soldiers in Virtual Environments: Route Learning and Transfer. U.S. Army Research Institute for the Behavioral and Social Sciences.

ATTACHMENT NUMBER 1

REPORTS AND REPORT DISTRIBUTION

REPORT TYPES

(a) Performance (Technical) Report(s) (Include letter report(s)) Frequency: Annual

(b) Final Technical Report, issued at completion of Grant.
NOTE: Technical Reports must have a SF-298 accompanying them.

(c) Final Financial Status Report (SF 269)

(d) Final Patent Report (DD 882)

REPORTS DISTRIBUTION		
ADDRESSEES	**REPORT TYPES**	**NUMBER OF COPIES**
Office of Naval Research Program Officer Terry Allard ONR 342CS Ballston Centre Tower One 800 North Quincy Street Arlington, VA 22217-5660	(a) & (b) w/(SF-298's)	3
Administrative Grants Officer OFFICE OF NAVAL RESEARCH REGIONAL OFFICE BOSTON 495 SUMMER STREET ROOM 103 BOSTON, MA 02210-2109	(c), (d) & SF-298's only for (a) & (b)	1
Director, Naval Research Laboratory Attn: Code 2627 4555 Overlook Drive Washington, DC 20375-5326	(a) & (b) w/(SF-298's)	1
Defense Technical Information Center 8725 John J. Kingman Road STE 0944 Ft. Belvoir, VA 22060-6218	(a) & (b) w/(SF-298's)	2
Office of Naval Research Attn: ONR 00CC1 Ballston Centre Tower One 800 North Quincy Street Arlington, VA 22217-5660	(d)	1

If the Program Officer directs, the Grantee shall make additional distribution of technical reports in accordance with a supplemental distribution list provided by the Program Officer. The supplemental distribution list shall not exceed 250 addresses.